REGIONAL CONVERGENCE: EVIDENCE FROM A NEW STATE-BY-STATE CAPITAL STOCK SERIES

Gasper A. Garofalo and Steven Yamarik*

Abstract—This paper seeks to reconcile the growth empirics technique of Mankiw, Romer, and Weil (1992) with the empirical results of Barro and Sala-i-Martin (1991) through the development of a new database covering the 1977–96 period. We create state-by-state capital stock and gross investment estimates by apportioning the national capital stock among the states. Using these estimates along with gross state product and employment data, we find evidence that the Solow growth model explains state-wide growth during this period. We consistently find a rate of convergence of around 2%. Our results, as a consequence, suggest that the empirical results of Barro and Sala-i-Martin are driven by the neoclassical growth process of Solow.

I. Introduction

The pattern and speed of regional income and convergence has been an issue for economists since the early 1960s.1 Recent studies by Barro and Sala-i-Martin (1991, 1992) find a 2% rate of convergence from the mid-1800s until the late 1980s. Carlino and Mills (1993) find similar results with a unit root test. Although the prediction of convergence is based on the neoclassical growth process, these studies do not directly estimate the neoclassical or Solow growth model. One framework available to directly test the Solow growth model is the growth empirics method of Mankiw, Romer, and Weil (1992) (hereafter, MRW). Although the MRW procedure was developed to investigate convergence at the international level, the technique has a natural application to U.S. regions and states because free factor mobility and a common technology base are more likely to exist among states than among nations. Unfortunately, the application of the MRW procedure to states has been hindered by the lack of a long-run and up-to-date database at the regional level. The only currently available database at the regional level was developed by Munnell (1990), and her data cover the period 1976–1986 and are difficult to update.2

This paper seeks to reconcile the growth empirics technique of Mankiw, Romer, and Weil with the empirical results of Barro and Sala-i-Martin through the development of a new database covering the 1977–1996 period. We create state-by-state capital stock and gross investment estimates by apportioning the national capital stock among the states. Using these estimates along with gross state product and employment data, we find evidence that the Solow growth model explains statewide growth during this period. We consistently find a rate of convergence of around 2%. Our results, as a consequence, suggest that the empirical results of Barro and Sala-i-Martin are driven by the neoclassical growth process of Solow.

The only test to date of convergence within the MRW model as it applies to U.S. states is by Holtz-Eakin (1993). He uses the Munnell data to estimate the growth-empirics model augmented with human capital. However, his estimation procedure differs fundamentally from MRW and ourselves. His test of conditional convergence is based on year-to-year variation in output per worker, which poses two limitations on his analysis: first, his estimation procedure employs inherently short-run data to a long-run model, and second, his estimation procedure does not allow him to parametrically estimate the speed of convergence. Because a confidence interval around his estimate is not given, one cannot determine if the speed of convergence is statistically different than Barro and Sala-i-Martin’s estimate.

Our database allows us to more extensively investigate the production structure and to estimate the growth-empirics model along the lines suggested by MRW. With this database, we test three aspects of regional convergence. First, we estimate the production structure of the states. A cross-section (48 states) time series (1977–1996) analysis shows strong evidence of constant returns to scale for a two-factor model (labor and capital) in which the estimated output elasticities match the observed income shares of two-thirds for labor and one-third for capital. When the model is augmented with human capital, we find that a portion of labor’s income share is payment for human capital. Second, we investigate the steady-state condition of the Solow growth model. Although we find evidence for steady-state convergence among states, the explanatory power of the model is modest.

Lastly, we consider the conditional convergence hypothesis. There is stronger statistical evidence for conditional convergence than for steady-state behavior. First, we find that each factor possesses an output elasticity equal to one-third. This result lends support for the rule-of-thumb estimate of MRW. Second, we find a rate of convergence that ranges from 1.5% (in a two-factor model) to 3.1% (in a model augmented with human capital). This range of values is consistent with the estimates found by Barro and Sala-i-Martin and Carlino and Mills.

The paper is organized as follows. We present the methodology used to estimate state-by-state capital stock and investment series in section II. We derive the Solow growth model along the lines of MRW in section III. We present the empirical results in section IV, and conclusions are presented in section V.

Received for publication January 26, 1999. Revision accepted for publication March 30, 2001.

* The University of Akron.

1 Early empirical work in this area includes Easterlin (1960) and Perloff et al. (1960). Borts and Stein (1964) provide a summary of the empirical evidence from 1860 to 1950 and a neoclassical framework to analyze the patterns.

2 Section II provides a description of Munnell’s technique along with a comparison to our method.
II. Data Methodology

Issues concerning regional productivity and growth can be investigated within a neoclassical framework only when a consistent and long-term capital stock series is available. Because the Bureau of Economic Analysis (BEA) (1998b) provides capital stock estimates for only the nation, the development of a state-by-state capital stock series is necessary. Our procedure is to apportion the national capital stock estimates to the states using annual income data. For each one-digit BEA industry, we apportion the national estimate by the relative income generated within each state. Each state’s capital stock estimate is then the sum of the industry estimates. We can represent our procedure by the following equations:

\[ k_{i,j}(t) = \frac{y_{i,j}(t)}{Y(t)} K_i(t) \]

\[ k_j(t) = \sum_{i=1}^{9} k_{i,j}(t) \]

where \( i \) represents the industry \((i = 1, \ldots, 9)\), \( j \) represents the state \((j = 1, \ldots, 51)\), the lowercase letters refer to amounts for state \( j \) and the uppercase letters refer to BEA totals for industry \( i \). Munnell (1990) also produces estimates of the private net capital stock for each state. Munnell apportions the BEA total for agriculture, manufacturing, and nonagriculture-nonmanufacturing using gross book value data from the economic censuses that occur every five years (1972, 1977, 1982, 1987). Her procedure decomposes the industrial total in the four census years using the relative gross book value within each state. In the intervening years, the capital stock in each industry is assumed to grow at the national rate.

The question remains whether our process of creating the capital stock series has produced any real information or simply replicated the relationship between national income and capital 51 times over. If our process replicated the national relationship, then the income-to-capital ratio across states should be close to the national ratio. Table 1 indicates that this is not the case. In 1947, the income-to-capital ratio was 0.46 for the nation. However, across states, the income-to-capital ratio in 1947 shows a good deal of variation: ranging from a low of 0.28 in West Virginia to a high of 0.59 in North Dakota. The standard deviation of the ratios is 0.68 in 1947. Moreover, there is also variability in the income-to-capital ratio through time. For example, table 1 shows that the rank order of the states changed significantly between 1947 and 1996. Also, the absolute value of the cumulative growth rate for 35 states was greater than the national growth rate.

Therefore, all cross-sectional variation in those years arises solely from cross-state differences in the industrial composition between agriculture, manufacturing, and nonagriculture-nonmanufacturing in the closest census year.

Our procedure improves on Munnell’s data in three ways. First, our procedure generates estimates for each year in the sample rather than for each census year. Second, our procedure improves on Munnell’s data in three ways. First, our procedure generates estimates for each year in the sample rather than for each census year. Second, our procedure improves on Munnell’s data in the sample rather than for each census year. As a result, cross-sectional variation in every year arises from cross-state differences in income-earned within each industry for that year. Second, our procedure improves on Munnell’s data in the sample rather than for each census year. As a result, cross-sectional variation in every year arises from cross-state differences in income-earned within each industry for that year. Third, our procedure improves on Munnell’s data in the sample rather than for each census year. As a result, cross-sectional variation in every year arises from cross-state differences in income-earned within each industry for that year.
Regression analysis confirms that our process simply did not replicate the national relationship between income and capital. A regression of the states' income-to-capital ratio on the national ratio for the 1947–1996 periods yields the following result:

\[
\ln \left[ \frac{y_j(t)}{k_j(t)} \right] = -0.1004 + 0.0042t + 0.8431 \ln \left[ \frac{Y(t)}{K(t)} \right] \\
(0.0341) \quad (0.00026) \quad (0.0609)
\]

\[\text{adj } R^2 = 0.0875.\]

An F-test on the hypothesis that the coefficient on \(\ln \left[ \frac{Y(t)}{K(t)} \right]\) equals 1 is rejected at the 1% level. However, if the relative income shares for each industry are added to the preceding regression, the point estimate on the national income-to-capital ratio changes very little to 0.8449, but the adjusted \(R^2\) rises to 0.4320. Therefore, more of the cross-state variation in the income-to-capital ratios can be attributed to variation in relative industry shares rather than the national income-to-capital ratio. As a consequence, there appears to be real information contained in our estimates of the private capital stock.
III. The Solow Growth Model

The Solow growth model starts with a constant returns-to-scale production function. There are two inputs, capital and labor, which are paid their marginal products. Assuming a Cobb-Douglas production function, output at time $t$ is

$$y_t = k_t^\alpha (A_t l_t)^{1-\alpha}$$

where $y_t$ is output, $A_t$ is the level of technology, $k_t$ is the stock of capital, and $l_t$ is the quantity of labor at time $t$.

The coefficients $\alpha$ and $(1-\alpha)$ represent the elasticities of output with respect to capital and labor, respectively. We assume that labor grows at the exogenous growth rate of $n$, that Hicks-neutral technical progress grows at the exogenous rate of $g$, and that capital depreciates at rate $\delta$.

Let $\hat{k} \equiv K/AL$ and $\hat{y} \equiv Y/AL$ represent the stock of capital and output in units of effective labor. Given the preceding assumptions and definitions, MRW show that $\hat{k}_t$ and $\hat{y}_t$ converge to the steady-state values $\hat{k}^*$ and $\hat{y}^*$ defined as

$$\hat{k}^* = \left[ \frac{s_k}{(n + g + \delta)} \right]^{1/(1-\alpha)}$$

and

$$\hat{y}^* = \left[ \frac{s_k}{(n + g + \delta)} \right]^{\alpha/(1-\alpha)}.$$

Substituting equation (4) into the production function in equation (3) yields the equation for the steady-state value of output per worker given by

$$\ln \left[ \frac{y_t}{l_t} \right] = \ln \left[ A_0 \right] + gt + \frac{\alpha}{1-\alpha} \ln \left[ s_k \right]$$

$$- \frac{\alpha}{1-\alpha} \ln \left[ n + g + \delta \right].$$

Equation (6) predicts that the log level of output per worker is positively related to the log difference between the gross investment rate, $s_k$, and the term $[n + g + \delta]$.

MRW find that augmenting the Solow model with measures of human capital improve its predictive power of explaining cross-country growth rates. We follow their second specification and incorporate human capital as a stock variable, $h_t$. As a result, equation (6) is modified as

$$\ln \left[ \frac{y_t}{l_t} \right] = \ln \left[ A_0 \right] + gt + \frac{\alpha}{1-\alpha} \ln \left[ s_k \right]$$

$$- \frac{\alpha}{1-\alpha} \ln \left[ n + g + \delta \right] + \frac{\beta}{1-\alpha} \hat{h}^*,$$

where $\beta$ represents the output elasticity on the stock of human capital and $\hat{h}^*$ represents the steady-state level of human capital per unit of effective labor.

The Solow growth model also makes predictions about the transition to a steady state. Approximating $\hat{y}$ around the steady state $\hat{y}^*$, the transition is given by $\partial \ln \left[ \hat{y}_t \right]/\partial t = \lambda (\ln \left[ \hat{y}^* \right] - \ln \left[ \hat{y}_t \right])$, where $\lambda = (n + g + \delta)(1 - \alpha)$. The parameter $\lambda$ is the speed of convergence. Solving this differential equation and then substituting equation (5) into equation (8) yields the testable implication:

$$\ln \left[ \frac{y_t}{l_t} \right] - \ln \left[ \frac{y_0}{l_0} \right] = (1 - e^{-\lambda t}) \frac{\alpha}{1-\alpha} \ln \left[ \frac{s_k}{s_0} \right]$$

$$- (1 - e^{-\lambda t}) \frac{\alpha}{1-\alpha} \ln \left[ n + g + \delta \right]$$

$$- (1 - e^{-\lambda t}) \ln \left[ \frac{y_0}{l_0} \right] + (1 - e^{-\lambda t}) \ln \left[ A_0 \right] + gt.$$
IV. Estimation Results

We estimate three aspects of the Solow growth model: the production structure (equation (3)), steady-state implications (equations (6) and (7)) and conditional convergence (equations (8) and (9)). These tests require data for output, labor, physical capital, human capital, and investment by state for the period 1977–1996.

A. Data

Our output measure is the real chained-weighted gross state product generated in the private sector. The data are reported by the Bureau of Economic Analysis (1998a). Our measure of labor is total employment on private payrolls reported by the Bureau of Labor Statistics (1998). Our human capital measure is the percentage of individuals, age 25 or older, who have completed four or more years of college. The data are taken from the 1980 Census of Population. Estimates of gross investment are derived from our capital stock series. The details of this procedure are presented in the appendix.

B. The Production Structure

A direct estimate of the production structure in equation (3) is not possible because we cannot observe the efficient units of labor. Therefore, we estimate a variant of equation (3) in which technical change is Hicks neutral. This form of the production function can be written as

\[ y_t = A_t k^{\alpha} l^{\eta}, \quad \alpha, \eta > 0, \tag{10} \]

where \( A_t = A_0 e^{\eta t} \). In logarithmic terms, the production structure is

\[ \ln[y_t] = \ln[A_0] + gt + \alpha \ln[k_t] + \eta \ln[l_t]. \tag{11} \]

We estimate equation (11) with pooled cross sections of contiguous states, and the time series covers 1977 to 1996. Ordinary least squares is appropriate only if the individual effects captured in the intercept term do not vary across states. This would imply that the level of multifactor productivity is the same among all states. We recognize that productivity measures are sensitive to differences in the industrial structures across states and to macroeconomic shocks that are not explicitly included in the model.

The ordinary least squares model can be generalized with a fixed-effects approach using the least squares dummy variable technique or with a random-effects approach using the variance components model. The fixed-effects approach allows the production structure to vary among the observations in the sample data in response to state-specific fixed effects and, as a result, takes into account within-state variations. The variance component technique is a special case of the random-coefficient model wherein only the intercept is treated as a random coefficient. The variance components model has the advantage of incorporating both within-state and between-state variations. Because we expect between-state variations to be fairly large, we prefer to estimate the production function with the error component technique. To determine the appropriateness of the random-effects model for our model, we employ a Hausman test suggested by Greene (1997). Applying the test to our data, we cannot reject the null hypothesis at the 5% level. As a consequence, we estimate equation (11) with a variance components model in which the cross-sectional units are the random component. The results are

\[ \ln[y_{i,t}] = 2.418 + 0.0073t + 0.3395 \ln[k_{i,t}] + 0.6682 \ln[l_{i,t}] \]

\[ (0.104) \quad (0.0003) \quad (0.0167) \quad (0.0186) \]

\[ \text{adj } R^2 = 0.9636. \]

The point estimates of the output elasticities sum to 1.0077, which is not significantly different from 1 at the 5% level. The results provide strong evidence for constant returns to scale. The point estimates, furthermore, are close to the standard rule of thumb wherein labor is paid two-thirds of national income. The rate of Hicks-neutral technical change is a rather modest 0.73% per year. This result implies that the majority of growth in output per worker can be explained by the growth in the capital-to-labor ratio. Because our results are consistent with the findings at the national level, we believe that our database provides the type of information necessary for empirical tests of regional growth and productivity.

We can augment the production structure by including a measure of human capital. Under the assumptions of Hicks-neutral technical change and constant returns to scale, we can express the production function with human capital as

\[ \ln[y_{i,t}] = \ln[A_0] + gt + \alpha \ln[k_t] + \beta \ln[h_{i,t}/l_t], \tag{12} \]

where \((h_{i,t}/l_t)\) is the stock of human capital embodied in each unit of labor at time \(t\). Our proxy for \((h_{i,t}/l_t)\) is the percentage of the population age 25 or older who have completed four or more years of college in 1980 (ED). This assumes that states with a higher value of ED have workers that embody additional...
a higher level of human capital. We again test the model using a variance component model in which the cross-sectional units are the random component. The results are

\[
\ln \left[ \frac{y_{i,t}}{l_{i,t}} \right] = 2.689 + 0.0074t + 0.3364 \ln \left[ k_{i,t}/l_{i,t} \right] + 0.1765 \ln [ED] \\
(0.105) (0.0002) (0.0166) (0.0641)
\]

\[ \text{adj } R^2 = 0.651. \]

The output elasticity on education is 0.1765. This implies that a 10% increase in the stock of education per worker will increase labor productivity by almost 2%. It is also interesting to note that the inclusion of ED does not significantly change the point estimate of \( \alpha \), and the return on capital remains at one-third.

C. Estimation of Steady-State and Convergence Equations

The steady-state and conditional-convergence equations are estimated under the assumption that the production structure is common to all states. This assumption implies that \( A_0 \) does not vary across states. In the Solow growth model, the steady-state and conditional-convergence equations are nonlinear in the parameters. Consequently, we use a nonlinear least squares estimator. One advantage of the nonlinear least squares estimator is that it gives us a confidence interval around the speed-of-convergence parameter.

Table 2 presents the estimation results of the steady-state and conditional-convergence equations. The asymptotic standard errors are reported in parentheses. For each equation, we ran a White (1980) test for the presence of heteroskedasticity. In each case, the hypothesis of a homoskedastic error structure could not be rejected at the 5% level. Thus, the standard errors need not be corrected for heteroskedasticity.

The left-hand side of table 2 reports the steady-state results. The data for the steady-state equations are cross-sectional where the data for each state is pooled into a single observation. Under both specifications, the estimate of the elasticity on physical capital, \( \alpha \), is positive and significant at the 1% level. Moreover, in the second column, the estimate of the elasticity on human capital, \( \beta \), is positive and significant at the 1% level. These results imply evidence of steady-state convergence among states from 1977 to 1996. If there is a common production structure among states with homogeneous inputs and freely available technology, the neoclassical adjustment process will ensure that all states converge to a common output-to-labor ratio. The clear policy implication is that attempts to attract capital and labor to a state will have only transitory effects. Regardless of economic development strategies, productivity levels among states will converge in the long run. Moreover, the human capital results suggest that 10% rise in ED will raise the output-to-labor ratio by almost 2% in the long run. This suggests that policies aimed at raising the percentage of the population with a college education will permanently increase the level of productivity in a state.

The results do require two notes of caution. First, the adjusted \( R^2 \) values are relatively low with values of 0.123 and 0.257, respectively. Second, the estimated output elasticities for capital, \( \alpha \), in both specifications are significantly below the values estimated in our production function. These two observations suggest that strong evidence for steady-state convergence is not warranted. The results, nevertheless, clearly point to a process of convergence at work among U.S. states.

\[ \text{The Hausman test for random effects yields an } m \text{ value of 0.2307, which is distributed chi-squared with two degrees of freedom. We cannot reject the null hypothesis at the 5\% level of confidence.} \]
The right side of table 2 presents the convergence results. First, we test for absolute or $\beta$-convergence. Introduced by Baumol (1986) and extended to U.S. states by Barro and Sala-i-Martin (1991), $\beta$-convergence involves the estimation of
\[
\ln \frac{y_t}{l_t} = \alpha + (1 - e^{-\lambda t}) \ln \left( \frac{y_0}{l_0} \right),
\]
where $\left( \frac{y_t}{l_t} \right)$ is the output-to-labor ratio in 1996, $\left( \frac{y_0}{l_0} \right)$ is the output-to-labor ratio in 1977, and $t = 20$.

The estimated rate of convergence ($\lambda = 0.021$) is consistent with the findings of Barro and Sala-i-Martin. Although the estimated value of $\lambda$ is significantly different from zero, the explanatory power of the model is relatively low with an adjusted $R^2$ equal to 0.148.

The last three columns in table 2 are tests of conditional convergence. The dependent variable is the cumulative growth rate. The values for $s_k$ and $(n + g + \delta)$ are mean values for 1977–1996, and the value for $h^*$ is the percentage of the population age 25 or older who have completed four or more years of college in 1980. The nonlinear estimation procedure requires us to assume that $(1 - e^{h^*}) \ln [A_0] + gt$ is constant across states. The assumption is necessary because we cannot directly observe the efficiency function $(A_t)$ by state.

The results for conditional convergence within the basic Solow model are more plausible than the steady state. First, the estimate of the output elasticity on $\alpha$ is 0.284, which is closer to the estimated value found in the production function and to the income share paid on capital. Second, the estimate on the rate of convergence is 0.025 and significant at the 1% level. The value is within the range found by Barro and Sala-i-Martin. However, the explanatory power of the model remains relatively low with an adjusted $R^2$ equal to 0.171.

The results for conditional convergence within the Solow model augmented with human capital are even more convincing. The estimates of $\beta$ are positive and significant at the 5% level with point estimates of 0.349 and 0.387 when regional dummies are included. The estimated impact of education under conditional convergence is significantly larger than under the production function and steady state. Moreover, given a point estimate of 0.310 for $\alpha$, the results provide strong support for the rule of thumb of MRW that labor, physical capital, and human capital each receive one-third share of output. The estimated rate of convergences are 0.031 and 0.017, which are well within the range found by Barro and Sala-i-Martin. Moreover, the inclusion
of dummy variables shows that the most-productive region is the northeast, which is excluded from the regression. The least-productive region is the west over our time period. Finally, the inclusion of the dummy variable considerably increases the explanatory power of the model by raising the adjusted $R^2$ to 0.465. We do recognize that, even with the inclusion of the regional dummies, the explanatory power of the model is relatively low. Nevertheless, we do believe that our results, taken as a whole, show a consistent pattern of conditional convergence among the U.S. states from 1977 to 1996.

V. Conclusion

We have developed a regional database that allows researchers to analyze important issues concerning regional output and productivity growth. Unlike currently available data sets, our methodology yields estimates that are current and easy to update. We apply this data to the issue of regional productivity convergence. We and easy to update. We apply this data to the issue of regional productivity convergence. We find that constant returns to scale is appropriate in a two-factor model yields results that are consistent with long-held beliefs. Nevertheless, we do believe that our results, taken as a whole, show a consistent pattern of conditional convergence among the U.S. states from 1977 to 1996.

REFERENCES


APPENDIX

This appendix describes the construction of a linear depreciation rate and a gross investment rate for each state. The BEA (1998b) provides data on the service life of capital equipment and the amount of capital equipment used in each industry annually for 1925. With a longer period from 1977 through 1996, the research can be extended to the post–World War II period. With a longer time period, the convergence of individual regions can be investigated. Furthermore, issues involving the productivity slowdown of the energy crisis and the effect of the extended growth period from the early 1980s to the late 1990s can be studied within a longer historical context.

This appendix describes the construction of a linear depreciation rate and a gross investment rate for each state. The BEA (1998b) provides data on the service life of capital equipment and the amount of capital equipment used in each industry annually for 1925–1996. Let $k$ be the index of the type of equipment, $e$, used. First, for each industry $i$ in state $j$, we construct a depreciation rate as a weighted average of the equipment used

$$
\delta_{i,j}(t) = \sum_{k=1}^{25} \left[ e_{i,j}(t)/k_{i,j}(t) \right] \cdot \left[ 1/\text{service life of } e_{i,j} \right].
$$

(A1)

Second, we derive a depreciation rate for each state $j$ as a weighted average of the capital stock in each industry $i$:

$$
\delta_{j}(t) = \sum_{i=1}^{9} \left[ k_{i,j}(t)/k_{i,j}(t) \right] \cdot \delta_{i,j}(t).
$$

(A2)

Note that equation (A2) generates a time series of linear depreciation rates for each state $j$. Using the data for equations (A1) and (1), we construct a gross investment series for each industry $i$ in state $j$

$$
i_{i,j}(t) = k_{i,j}(t) - \left[ 1 - \delta_{i,j}(t) \right] k_{i,j}(t - 1).
$$

(A3)

Lastly, equation (A3) is summed across industries to arrive at a gross investment series for each state $j$

$$
i_{j}(t) = \sum_{i=1}^{9} i_{i,j}(t).
$$

(A4)

The estimates of equations (A2) and (A4) are used to construct the depreciation and investment rates used in the steady-state and convergence tests of the Solow growth model.